DOCUMENT RESUME

ED 074 712 EM 010 851

AUTHOR Sulzen, Robert H.
TITLE The Effects of Empirical Program Revision and the Presentation of Objectives on Student Performance.
PUB DATE 27 Feb 73
NOTE 14p.; Paper presented at the American Educational Research Association Annual Meeting (New Orleans, Louisiana, February 27, 1973)

EDRS PRICE MF-$0.65 HC-$3.29
DESCRIPTORS *Behavioral Objectives; Educational Improvement; Educational Objectives; *Educational Research; Effective Teaching; *Instructional Design; *Instructional Technology; Military Training; Performance; *Performance Factors; Training Objectives

ABSTRACT

The two techniques for improving the quality of programed instruction, empirical revision, and the presentation of objectives before instruction were tested by this experiment in order to obtain a systematic replication of former experimental results. This study was designed to investigate: whether the presentation of the specific objectives technique consistently resulted in improved student performance; whether empirical revision resulted in improved student performance across subject matter areas and student populations; and whether a combination of both techniques would have a greater effect than either one alone. Using a 2x2 research design employing placebo objectives and using student performance as the dependent variable, the study showed that the combination revision/objective treatment was statistically superior to either technique alone. The experiment was conducted on Officer Candidate Personnel at the United States Army Infantry School. (MC)

ED 074712

NPO
EM

THE EFFECTS OF EMPIRICAL PROGRAM REVISION AND THE PRESENTATION OF OBJECTIVES ON STUDENT PERFORMANCE

by

Robert H. Sulzen*

PRESENTATION GIVEN AT THE ANNUAL
MEETING OF THE AMERICAN
EDUCATIONAL RESEARCH ASSOCIATION
NEW ORLEANS, LOUISIANA
27 February 1973

U.S. DEPARTMENT OF HEALTH,
EDUCATION & WELFARE
OFFICE OF EDUCATION
THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL OFFICE OF EDUCATION POSITION OR POLICY

EM 010851

*The opinions or assertions contained herein are the private ones of the author and are not to be construed as official or reflecting the views of the Department of the Army or the Army at large. At the time the research was conducted the author was employed by the Department of the Army. He is presently with the Training Technology Branch, Advanced Systems Division, Air Force Human Resources Laboratory.

The development of high quality instructional programs has been a persistent problem for instructional technology. High quality is dependent upon student performance following instruction. Two techniques for improving the quality of programmed instruction have demonstrated their effectiveness in experimental studies.

The first technique (Silberman, et al., 1964) is the revision of the instructional program by empirical tryout with revisions based on student difficulties and comments. This procedure is followed by subsequent trials until the satisfactory level of student performance is attained.

The second technique is the presentation of specific objectives before receiving instruction (Mager and McCann, 1961). This technique is far less costly than the first and may be applied with far less trouble.

The presentation of specific objectives to students prior to learning is a relatively simple modification to a program. It requires only that the objective be stated in a way that will contribute to the desired change when presented to the students. Empirical revision on the other hand, is a complex and involved process requiring the extended time of a revisor and the time of trial subjects plus the costs of making the necessary revisions. Additionally, the process of revising a program on the basis of student performance often appears to be peculiar to the individual conducting it (Markle, 1967b, p. 124). From the same data it is quite probable that different revisors will follow very different plans of program improvement. The implication is clear; if presenting the objectives before the program proves as effective as empirical revision, the cost and effort is considerably less and the former strategy should be employed.

There is, of course, a key question that may be asked of the two major treatments cited (objectives before instruction and empirical revision): do these interventions produce reliable results? A particular experimental treatment may not be significant in reality, but if it is tested repeatedly it is likely to appear significant in a few attempts (Sidman, 1960, p. 45). The few attempts that result in an estimate of significance are the experiments that are most likely to be published, particularly if editors continue to be more sensitive to significant results than scientific rigor (Lykken, 1967, p. 159). A replication of the results supporting the two major treatments in this study under different circumstances would lend credence to their generalizibility.

If a particular experimental treatment is sufficiently powerful, it should give relatively comparable results when other factors in the environment vary. Many studies in educational research consist of short treatments speedily analyzed and reported (Shulman, 1970, p. 391). Experimenters should attempt systematic replication (Sidman, 1960, p. 111) where the findings in question hold up under varying conditions, such as student populations, subject matter areas, and programmers (or revisors).

The present study was designed to investigate the following questions: 1) Does the presentation of specific objectives prior to programmed instruction consistently result in improved student performance across subject matter areas and student populations? 2) Does empirical revision result in improved student performance across subject matter areas, student populations, and revisions? 3) If both techniques are employed together, does one result greater levels of student performance than the other?

The dependent variable was student performance on the criterion test for the specific instructional sequence utilized. The independent variables were the type of objective given to students prior to learning (specific or placebo) and the revision status of the program (revised or unrevised). The placebo objectives were content and goal relevant information lacking the specificity of the other treatment. The unrevised program was instruction that did not bring about the desired level of student achievement. The empirically revised program was the modification made to the unrevised program by one revisor. In order to compare the unrevised program by one revisor. In order to compare the effects of the major variables and to test any interaction a 2 x 2 research design was the basis for analysis (See Figure 1).

		OBJECTIVES	
		Placebo Objective	Specific Objective
REVISION	Unrevised Program	Control C	Objective O
	Revised Program	Revised R	Revised/Objective RO

Figure 1. 2 x 2 Research Design

From Figure 1 it may be seen that there were four treatment groups as follows:

(C) Control--Placebo objectives and unrevised program
(O) Objective--Specific objectives and unrevised program
(R) Revised--Placebo objectives and revised program
(RO) Revised/Objective--Specific objectives and revised program

Since the study involved nine separate replications, the final analysis was made with a 2 x 2 x 9 factorial design.

The null hypotheses for these tests were as follows:

1. There will be no significant differences on the criterion test scores between the treatment groups given empirically revised programs (R,RO) and the treatment groups given the unrevised programs (C,O).

2. There will be no significant differences on the criterion test scores between the treatment groups given specific objectives prior to the programs (O,RO) and the treatment groups given a placebo objective prior to instruction (C,R).

3. There will be no significant interaction among the major treatments of revision, objectives, or replications.

No hypotheses were formulated concerning the different replications as a main effect. Since different criterion instruments were utilized for different programs, it was necessary to equate student performance across replications. This was accomplished by converting the dependent variable scores to z-scores; therefore, the equated groups should not differ significantly from each other. The possibility still existed that there could be significant interactions between the replications and the objective and/or revision variables; therefore, the null hypotheses concerning interactions of the replications was asserted.

METHOD

Revision. Two empirical revision sessions were conducted to revise existing programs. The first session was two weeks and the second three weeks. The general procedures for the revision process are described elsewhere (Sulzen, 1970). The final revision session (which provided six of the revisions) incorporated the following specific practices: (1) a personnel selection technique (Lysaught and Pierleoni, 1970) to obtain effective revisors, (2) an opportunity for revisors to try out materials on four individual students (developmental trials) and on four groups, and (3) replicable instruction in the form of nine selected lessons from a tape and slide program (Deterline, 1970).

Programs. Programs (Table I) varied in format and length. All programs were in tape and slide format except for one programmed text. The tape and slide programs had questions interspersed throughout the instruction. These were multiple choice questions for eventual use in a student response classroom using the EDEX system. This system allows students to respond to a question by selecting an alternate at his desk on a responder. The responses are displayed immediately for instructor information and automatically recorded for later analysis. The length of time for each program varied between half an hour and two hours (in no case did the revised program exceed the unrevised program in length).

Criterion Tests. Table I shows the criterion "responses" for each program. The term "response" was used instead of item because a one-to-one relationship did not exist. Some items required more than one response. Scores reflected a weighting system in use by the Instructional Department of the Infantry School with responsibility for that particular subject matter.

Three types of student responses were required in the criterion tests. These responses were selecting the proper alternative for multiple choice questions, matching a term or number to an appropriate description for matching questions, and writing a short answer for competion questions. The short answer was usually no more than two or three words.

Subjects. The subjects in the experiment were members of the U.S. Army undergoing leadership training at the U.S. Army Infantry School Fort Benning, Georgia, during the summer of 1970. They were either Officer Candidates or Noncommissioned Officer Candidates. They were assigned to companies (classes) based upon when they arrived at the school. Classes were utilized for the experiment. Selection of each class was based upon its possessing no prior training on the subject matter presented and upon administrative constraints. All leadership candidates at the Infantry School were in a stressful situation during the time period of the study. The typical morning began at 5:30 a.m. with reveille, breakfast at 6:00 a.m., an inspection of quarters at 7:00 a.m., and company formation to go to class at 7:30 a.m.. Each day consisted of 8 to 12 hours of training either in a classroom, on a firing range, or in a field exercise. The ages of the candidates were between 18 and 24. Although some candidates were older, they were eliminated from the study because they typically had received prior training on the subject matter presented. Another cause of experimental mortality was the assignment of the duty Charge-of-Quarters (CQ). These candidates were required to stay awake most of the night and still participate in the training the next day.

Staff. Members of the Instructional Systems Branch at the U.S. Army Infantry School served as the experimental staff. The staff insured that all companies (classes) used in the experimental replications were randomly assigned to one of the four major treatment groups. They presented the objectives given prior to the instruction in the prescribed manner. They were present during the instruction to detect the presence of extraneous factors. They administered the immediate criterion test following the instruction, and they scored the criterion tests following the replication.

Setting. The U.S. Army Infantry Hall was the site for all tests and consisted of a large building erected in the mid-1960s. There were several classrooms of two sizes: 200-man and 50-man. The building had a thermostatically controlled environment at 75 degrees.

TABLE I

INFORMATION ON EXPERIMENTAL REPLICATIONS

Replication	Subjects[1]	Program Title[2]	Program Length[5]	Criterion Responses	Time of Day[6]
1	OC	Electronic Warfare	1/2	21	Afternoon
2	OC	Evasionn and Escape[2]	1	20	Afternoon
3	OC	Supply Procedures[3]	2	34	Morning
4	NCOC	Patrolling Terms	1	44	Evening
5	NCOC	Fundamentals of Defense[4]	2	49	Evening
6	OC	Fundamentals of Defense[4]	2	51	Morning
7	OC	Supply Procedures[3]	2	34	Morning
8	OC	Maintenance Records	2	35	Afternoon
9	OC	Evasion and Escape[2]	1	35	Afternoon

[1]OC = Officer Candidate; NCOC = Noncommissioned Officer Candidate.

[2]Revised twice because of a change in objectives.

[3]Revised twice because of a poor first revision.

[4]Revised by different revisors; one for each target population.

[5]Approximate time in hours.

[6]General time of program and test administration.

Treatments. There were four treatments: 1) the first was the control group (C) which received the unrevised instructional program and a placebo objective before the instruction; 2) the second group was the objective group (O) which received the same unrevised instructional program, but was given a specific objective prior to the program; 3) the third group was the revision group (R) which received an empirically revised program and a placebo objective; 4) the fourth group was the revision and objective group (RO) which received both the revised program and a specific objective.

Procedure. A company (class) was administratively assigned to participate in the experiment. Individual students were randomly assigned to one of the four treatment groups. The two groups assigned to receive the unrevised instruction (C,O) went to one classroom while the groups assigned to receive the revised instruction (R,RO) went to a second classroom. Inside the classroom the group assigned the specific objective (O,RO) sat on one side, while the group assigned the placebo objective (C,R) sat on the other side. The appropriate objectives were placed face down on the tables. A member of the experimental staff then read the following statement:

LISTEN CAREFULLY TO THESE INSTRUCTIONS:

1. Keep the sheet of paper on your desks face down.

2. On this paper you will find the objectives for the instruction you are about to receive.

3. WHEN you are told to do so, read the sheet carefully. It should help you to answer questions on the examination you will receive later.

4. You may turn the objective sheet over now.

5. Put your name and roster number on the top right corner.

(Pause a few seconds)

6. Read the objectives carefully. Try to remember them.

The students were then allowed to read the objectives. The objectives were collected when all the students had finished. The sheets with objectives having both names and roster numbers were used to verify test results with assigned treatment.

The students were then presented either the unrevised program (C,O) or the revised program (R,RO). Immediately following the program the subjects were given the criterion test. The tests were scored and then each individual was identified by treatment group. This assignment was verified by checking the student name and roster number written on the

objective sheets. The class leaders were consulted and any student assigned late night duties was eliminated from the experiment. Scores were tabulated and submitted for analysis.

RESULTS

The dependent variable for each replication in the major experiment was statistically equated. This was accomplished by converting all the scores in one replication to z-scores (based on the performance of all subjects in the replication). Once replications with different dependent variables were equated by the z-scores, a 2 x 2 x 9 analysis of variance was made using the z-scores. The summary of this analysis (shown in Table II) indicated that the main effects of both revision and objectives were significant at the .01 level. Replications could not be significant

TABLE II

ANALYSIS OF VARIANCE SUMMARY TABLE FOR THE EFFECT OF PRESENTING SPECIFIC OBJECTIVES BEFORE INSTRUCTION AND OF EMPIRICAL REVISION ON THE CRITERION BEHAVIOR OF MILITARY STUDENTS

Source	SS	df	MS	F
Revision	38.81	1	38.81	47.43*
Objectives	5.87	1	5.87	7.18*
Replications	.24	8	.03	.04
Revision X Objectives	.08	1	.08	.10
Revision X Replications	22.70	8	2.84	3.47*
Objectives X Replications	5.54	8	.69	.85
Three Factor Interactions	8.99	8	1.12	1.37
Experimental Error	256.13	313	.82	

*Significant at the .01 level

since they were equated by z-score conversion for each replication. The only interaction that was significant was the interaction between revision and replications. This interaction was the result of the differential success enjoyed by the revisors.

The most appropriate post hoc analysis was considered to be a comparison of each of the four major treatments (control, objectives, revision and revision/objectives). A Newman-Keuls method was employed to make the comparison which is shown in Table III. The results indicate that all other treatments were significantly better in terms of student performance than the control. Empirical revision was significantly better

TABLE III

NEWMAN-KEULS TESTS ON THE DIFFERENCES BETWEEN THE FOUR MAJOR TREATMENT MEANS

Treatments		Control	Objective	Revision	Revision/Obj
	Means	-.476	-.188	.222	.441
Control	-.476	-	.288*	.698**	.917**
Objectives	-.188		-	.419**	.629**
Revision	.222			-	.219
Revision/ Obj	.441				-

		r=2	r=3	r=4
	$q.99\ (r,120)$	3.70	4.20	4.50
	$q.95\ (r,120)$	2.80	3.36	3.69
$\sqrt{MS_{e/\tilde{n}}}$	$q.99\ (r,120)$	.355	.403	.432
$\sqrt{MS_{e/\tilde{n}}}$	$q.95\ (r,120)$	.269	.323	.354

*Significant at the .05 level
**Significant at the .01 level

than presenting objectives prior to instruction, but utilization of both objectives prior to instruction and empirical revision was not significantly better than empirical revision alone. Nevertheless, it is evident that the revision/objective treatment group was numerically superior to the revision group and the results approach significance. Inspection of

the scores for different replications reveals that the revision/objective group was consistently high or the highest. It is possible that a directional hypothesis (H_1: RO>R>O>C) might hold with a different statistical technique (see Jonckheere, 1954) but this type of relationship was not predicted in this study.

DISCUSSION

It may be concluded that each of the null hypotheses may be rejected with a result significant at the .01 level. In the case of the interaction hypotheses, only that portion dealing with the interaction of revision and replications may be rejected. A Newman-Keuls post hoc analysis of the four treatment variables indicates the following relationship: RO = R>O>C.

The answer to the research questions are as follows: The presentation of objectives prior to programmed instruction was found to be a significant variable as measured by student performance replicated across student populations and subject matter areas. Empirical revision was found to significantly improve student performance across subject matter areas, student populations, and revisions. When both empirical revision and the presentation of objectives are employed together, revision results in significantly higher student performance. This last finding must be viewed in light of the fact that empirical revision required 80 to 120 man-hours per program, as opposed to approximately one man-hour per program for writing the specific objectives.

The results of this study reaffirm by replication the effectiveness of empirical revision and the presentation of specific objectives to learners prior to programmed instruction. The fact that these experimental variables proved effective under the systematic replication of this study (where the conditions of subject matter areas, student population, and revisor were varied) lends even stronger credence to their generalizability. It is recommended that these practices be continued by educational technologists when developing instructional materials.

Empirical revision is a powerful technique for improving student performance following programmed instruction, but it is dependent upon the effectiveness of the revisor. In the study conducted, a technique for selecting instructional programmers with standard psychological tests (Lysaught and Pierleoni, 1970) was not entirely accurate in predicting successful revision, but the revisors selected by this method were not volunteers. Both the experimenter and the Instructional Systems Branch were of the opinion that the selection procedure predicted those capable of successful revision, but could not predict those motivated to do so. It is recommended that those using the Lysaught-Pierleoni selection technique draw programmers or revisors only from volunteers.

Although empirical revision appears to be more effective than the presentation of objectives prior to learning, the empirical revision process is far more costly in terms of man-hours than the objectives technique (80 to 1 and 120 to 1). When cost is an important factor and the student performance is close to the desired level, it is recommended that objectives be presented to the learner prior to instruction instead of undertaking a complete empirical revision. When time is an importnat factor and the student performance is close to the desired level, it is recommended that objectives technique be employed. When student performance is well below the desired achievement level, empirical revision is the recommended approach.

References

Ausubel, David P. "The Use of Advanced Organizers in the Learning and Retention of Meaningful Verbal Material", Journal of Educational Psychology, 51:5 (October 1960) 267-272.

Baker, Eva L. "The Generalizability of Rules for Revision of Programs", Paper presented at the Annual Convention of the National Society for Programmed Instruction, Anaheim, California, April-May 1970.

Campbell, Donald T. and Julian C. Stanley. "Experimental and Quasi-Experimental Designs for Research on Teaching", Handbook of Research on Teaching. Edited by N. L. Gage. Chicago: Rand McNally and Company, 1963, 171-246.

Deterline, William A. Principles and Practice of Instructional Technology. (A Filmstrip and Audio Tape Program) Palo Alto, California: General Programmed Teaching, 1970.

Gilbert, Thomas F. "Mathetics: II. The Design of Teaching Exercises", Journal of Mathetics, 1 (April 1962) 7-56.

Gilbert, Thomas F. "On the Relevance of Laboratory Investigation of Learning to Self-Instructional Programming", Paper presented at the 1959 convention of the American Psychological Association at Cincinnati, Ohio. Also in Teaching Machines and Programmed Learning. Edited by Robert Glaser and Arthur A. Lumsdaine. Washington, D.C.: National Education Association of the United States, 1960, 475-485.

Groper, George L., Arthur A. Lumsdaine, and Virginia Shipman. Studies in Televised Instruction, Report Number 1: Improvement of Televised Instruction Based on Student Responses to Achievement Tests. AIR-C13-61-FR-245(I). Pittsburgh, Pennsylvania: American Institute for Research, 1961.

Harrison, Grant V. "Instructional Value of Presenting Explicit vs. Vague Objectives", Paper presented at the Annual Convention of the California Educational Research Association, March 1967.

Horton, George A., Jr. Personal communication on the results of doctoral research conducted while a doctoral candidate at the University of California, Los Angeles, 1969.

Jenkins, Joseph R. and Stanley L. Deno. "Effects of Instructional Objectives on Learning", Paper presented at the Annual Convention of the American Educational Research Association, Minneapolis, Minnesota, March 1970.

Jonckheere, A. R. "A Distribution-Free k-Sample Test Against Ordered Alternatives", Bionetrika, 41 (1954) 133-145.

Lykken, David T. "Statistical Significance in Psychological Research", Psychological Bulletin, 70:3 (1967) 151-159.

Lysaught, Jerome P. and Robert G. Pierleoni. "Predicting Individual Success in Programing Self-Instructional Materials", AV Communication Review, 18:1 (Spring 1970) 5-24.

Mager, Robert F. Preparing Objectives for Programmed Instruction. San Francisco: Fearson Publishers, Inc., 1962.

Mager, Robert F. and J. McCann. Learner-Controlled Instruction. Palo Alto, California: Varian Associates, 1961.

Markle, David G. "Controlling Behavior Changers' Behavior" AV Communication Review, 16:2 (Summer 1968a) 188-203.

Markle, David G. Final Report: The Development of the Bell System First Aid and Personal Safety Course. AIR-E81-4/67-FR, Palo Alto, California: American Institutes for Research, 1967a.

Markle, David G. Programmed Instruction: Developmental Process. (A Film) Distributed by: National Audio Visual Center, National Archives and Records Service, Washington, D.C., 1968b.

Markle, Susan M. "Empirical Testing of Programs". Programed Instruction. The Sixty-Sixth Yearbook of the National Society for the Study of Education. Edited by Phil C. Lange. Chicago: The University of Chicago Press, 1967b, 104-138.

Markle, Susan M. Good Frames and Bad: A Grammer of Frame Writing. New York: John Wiley and Sons, 1964.

Miles, David T. et al. "The Effects of Study Questions on College Students' Test Performances", Psychology in the Schools, (January 1967) 25-26.

Robeck, Martin J. "A Study of the Revision Process in Programmed Instruction", Unpublished Master's Thesis, University of California, Los Angeles, 1965.

Rosen, Marvin J. "An Experimental Design for Comparing the Effects of Instructional Media Programing Procedures", Doctoral Dissertation, University of California, Los Angeles, 1968.

Shulman, Lee S. "Reconstruction of Educational Research", Review of Educational Research, 40:3 (June 1970) 371-396.

Sidman, Murray. Tactics of Scientific Research. New York: Basic Books, Inc., 1960.

Silberman, Harry F., John Coulson, Ralph Melaragno, and Gerald Newmark. Use of Exploratory Research and Individual Tutoring Techniques for the Development of Programed Methods and Theory. Santa Monica, California: Systems Development Corporation, 1964.

Sulzen, Robert H. "The Formulation of a Replicable Means of Empirical Development or Revision of Automated Programmed Instruction", a presentation given at the Annual Convention of the National Society for Programmed Instruction, Anaheim, California, May, 1970.

Tiemann, Phillip W. "Student Use of Behaviorally-Stated Objectives to Augment Conventional and Programmed Revisions of Televised College Economics Lectures", Paper presented at the Annual Convention of the American Educational Research Association, February, 1968.

Veldman, Donald J. Fortran Programming for the Behavioral Sciences, New York: Holt, Rinehart and Winston, 1967.

Verhaegen, Ritchie M. "The Effect of Learner-Controlled Instruction in a Tenth Grade Biology Curriculum", Unpublished Master's Thesis, University of California, Los Angeles, 1964.

Weisburg, Joseph S. "The Use of Visual Advance Organizers for Learning Earth Science Concepts", Journal of Research in Science Teaching, 7:2 (1970) 161-165.

Winer, B. J. Statistical Principles in Experimental Design. New York: McGraw Hill, Inc. 1962.